Une illustration du rapport sur la géométrie de la commission Kahane : l'exemple de la géométrie affine du collège

Daniel PERRIN

0. Introduction.

Le texte qui suit a pour but d'expliciter, sur l'exemple de la géométrie affine du collège (Thalès, médianes, etc.), certaines des conceptions mathématiques et épistémologiques qui sous-tendent le rapport d'étape de la commission de réflexion sur l'enseignement des mathématiques (commission Kahane).

L'un des points de départ de notre réflexion est la citation suivante de Bourbaki (cf. [Bo] p. 142) sur la géométrie "élémentaire" :

Mais la situation devient bien plus nette avec les progrès de la théorie des invariants [...] qui marque la mort, comme champ de recherches, de la théorie classique des invariants elle-même et de la géométrie "élémentaire", qui en est devenue pratiquement un simple dictionnaire. Sans doute, rien ne permet de prévoir a priori, parmi l'infinité de théorèmes que l'on peut ainsi dérouler à volonté, quels seront ceux dont l'énoncé, dans un langage géométrique approprié, aura une simplicité et une élégance comparables aux résultats classiques, et il reste là un domaine restreint où continuent à s'exercer avec bonheur de nombreux amateurs (géométrie du triangle, du tétraèdre, des courbes et surfaces algébriques de bas degré, etc.) Mais pour le mathématicien professionnel, la mine est tarie ...

Cette phrase est intéressante à un double titre, au point de vue mathématique d'abord, mais aussi à cause des conséquences que cette position a pu avoir sur l'enseignement de la géométrie. En effet, cette prise de position de Bourbaki, que l'on peut considérer comme emblématique des mathématiciens de l'époque, a joué un rôle important dans la réforme des mathématiques modernes à la fin des années 60. Le discours était en gros le suivant : puisque la géométrie élémentaire est morte, mieux vaut enseigner autre chose, les espaces vectoriels, les transformations, l'analyse,... On relira à ce sujet l'introduction de Jean Dieudonné dans [D].
Bien sûr, avec le recul, tout le monde s'accorde à dire que la réforme des mathématiques modernes a été un échec. Bien sûr, on y trouve facilement de multiples raisons du côté de la formation des maîtres ainsi que de nombreuses raisons didactiques. Ces raisons, pour importantes qu'elles soient, ne nous ont pas paru totalement satisfaisantes et nous avons cherché à comprendre s'il n'y avait pas aussi de bonnes raisons mathématiques et épistémologiques à cet échec de la réforme. Ces idées ont été développées dans la deuxième partie du rapport en

une analyse de la réforme des mathématiques modernes qui s'articule autour de la critique de cinq points :

- *Le projet du "tout linéaire".*
- *La minoration systématique du rôle des figures.*
- *La minoration du rôle des invariants, notamment aire et angle.*
- *L'abandon des cas d'égalité des triangles.*
- *La disparition des géométries riches.*

Nous allons parler ici du troisième point (le rôle des invariants, et notamment de l'aire), en reprenant avec plus de détails le contenu de l'annexe 1 du rapport. Si nous avons choisi la notion d'aire c'est parce qu'elle est un peu la mal aimée de notre enseignement. Elle allie pourtant de nombreux avantages : elle est très intuitive, c'est un excellent outil de démonstration et enfin, et c'est ce que nous voudrions montrer ici, son utilisation est tout à fait justifiée par la théorie. Une autre raison de ce choix c'est que la théorie, dans le cas de la géométrie affine du triangle, est particulièrement simple. Bien entendu, on peut mener une réflexion assez voisine dans le cas de la géométrie euclidienne avec les invariants longueur et angle.

Attention, le plaidoyer développé ci-dessous en faveur de la notion d'aire ne signifie évidemment pas que les autres outils disponibles actuellement au collège, et notamment les transformations, aient perdu de leur importance. L'objectif à poursuivre est plutôt de trouver un nouvel équilibre entre divers points de vue.

1. Le programme d'Erlangen et les invariants.

a) Le programme d'Erlangen.

Il s'agit de la thèse de Felix Klein soutenue à Erlangen en 1872. C'est un texte qui a inspiré nombre de promoteurs de la réforme des mathématiques modernes.
Le travail de Klein se veut une unification de toutes les géométries que le dix-neuvième siècle a vu éclore, à côté de la géométrie euclidienne classique : géométries projective, anallagmatique (la géométrie de l'inversion), non euclidiennes, etc. Sa thèse est qu'une géométrie consiste, pour l'essentiel, en la donnée d'un ensemble X et d'un groupe G de transformations de X. Les éléments de G sont les transformations "permises" dans la géométrie en question et ils caractérisent cette géométrie. Il s'agit, par exemple, des transformations affines pour la géométrie affine plane, ou des isométries affines pour la géométrie euclidienne plane ou encore des homographies pour la géométrie projective. Les propriétés relatives à la géométrie en question (propriétés affines, euclidiennes, projectives) sont celles qui sont conservées dans l'action du groupe, ainsi que le dit Klein (cf. [K] p.7) :
"*Étant donnés une multiplicité et un groupe de transformations de cette multiplicité, en étudier les êtres au point de vue des propriétés qui ne sont pas altérées par les transformations du groupe.*"
Par exemple, pour citer trois résultats célèbres, Pappus, qui n'emploie que les notions de concourance et d'alignement, est un théorème projectif tandis que Thalès, qui utilise des parallèles, est un résultat affine et Pythagore, qui met en jeu longueurs et orthogonalité, est un théorème euclidien. On peut dire, en

quelque sorte, que chaque théorème possède une niche écologique privilégiée, qui correspond au cadre dans lequel il s'énonce avec le plus de généralité.

b) Erlangen et les invariants.

Lorsqu'on lit Klein avec les yeux d'un mathématicien actuel on ne peut qu'être en accord avec ce qu'il écrit, mais on reste un peu sur sa faim : si le groupe des transformations permises indique bien dans quel type de géométrie on travaille, il ne dit pas comment obtenir les théorèmes de cette géométrie. En fait, ce problème est résolu au moyen de la théorie des invariants (à peu près contemporaine de Klein). C'est à cela que Bourbaki fait allusion. Notre opinion est que cette théorie est inséparable du programme d'Erlangen et, pour utiliser une image familière, que le programme d'Erlangen sans la théorie des invariants c'est comme un vélo sans pédales. D'ailleurs, Klein lui-même assigne à la géométrie l' objectif suivant : *on donne une multiplicité et un groupe de transformations de cette multiplicité ; développer la théorie des invariants relatifs à ce groupe.*

c) Les invariants, version naïve.

Considérons d'abord les invariants en un sens intuitif. Il s'agit simplement alors des notions qui sont conservées par les transformations du groupe en question. Par exemple, dans le cas de la géométrie euclidienne les invariants les plus immédiats sont les notions usuelles de longueur, d'orthogonalité ou plus généralement d'angle.

Dans le cas de la géométrie affine qui va surtout nous intéresser ici, longueur et d'angle ne sont plus des invariants, comme on le voit sur l'exemple de l'affinité ([1]). On dispose cependant d'un semi-invariant qui est l'aire, par exemple celle des triangles. En effet, même si on définit souvent l'aire à partir d'une structure euclidienne (penser à la formule $base \times hauteur/\,2$), l'aire est en réalité une notion affine.([2]) Nous le verrons abondamment dans ce qui suit, mais il suffit pour s'en convaincre de penser à la transformation qui consiste à déplacer le sommet d'un triangle sur une parallèle à la base, transformation qui n'est certes pas une isométrie, mais qui conserve pourtant l'aire.
Nous disons que l'aire est un semi-invariant seulement car si u est une transformation affine quelconque (penser à une affinité ou une homothétie), l'aire n'est pas conservée, mais multipliée par le déterminant de l'application linéaire associée comme nous le verrons au paragraphe 3 ci-dessous. Toutefois c'est un vrai invariant pour les transformations affines de déterminant 1 et, même dans le cas des transformations affines générales, les rapports d'aires, eux, sont conservés.([3])

d) Invariants naïfs et théorèmes.

L'une des critiques essentielles que l'on peut faire à la réforme des mathématiques modernes est d'avoir minoré l'importance des invariants. Dans le cas de la géométrie euclidienne (resp. affine) cela a conduit à négliger un peu trop la

([1]) Rappelons que l'affinité (ou dilatation) d'axe Ox, de direction Oy et de rapport λ, est la transformation qui au point de coordonnées (x, y) associe $(x, \lambda y)$.

([2]) D'ailleurs, toute la construction d'une mesure des aires telle que la pratique par exemple Henri Lebesgue dans [L] peut être faite sans la structure euclidienne en prenant un parallélogramme comme unité au lieu d'un carré.

([3]) On voit apparaître ici les rapports de grandeurs chers aux grecs.

notion d'angle (resp. la notion d'aire), au profit d'un usage quasi exclusif des transformations.
Nous allons essayer de montrer sur l'exemple de la géométrie affine que l'usage des invariants est pourtant une méthode parfaitement pertinente en montrant au paragraphe 2 comment un petit nombre de lemmes qu'on peut enseigner au collège permettent de prouver une grande quantité de résultats de géométrie affine. Cette constatation empirique, que nombre de professeurs ont d'ailleurs faite avant nous, sera justifiée par la théorie au paragraphe 3 où nous expliquerons que l'aire est, en fait, le seul invariant de la géométrie affine du triangle, et que tous les théorèmes de géométrie affine (et c'est aussi le cas des autres géométries) découlent de la connaissance des invariants et des relations qui les lient.

2. L'invariant aire et la géométrie du collège.

Dans ce paragraphe nous travaillerons dans un plan affine (que l'on pourra supposer euclidien, encore que cette dernière condition ne soit pas vraiment nécessaire). Le lecteur pourra, au choix, penser ce plan comme défini à l'aide d'une axiomatique du type de celle d'Euclide (points, droites, parallèles, etc.) ou à l'aide de la notion d'espace vectoriel. On sait depuis Hilbert que ces deux approches sont, en fait, équivalentes. Nous noterons les points avec des lettres minuscules, préférant réserver l'usage des majuscules aux parties du plan. Le lecteur nous pardonnera cette entorse à l'usage de l'enseignement secondaire, car elle lui facilitera peut-être la lecture du paragraphe 3.

Nous aurons évidemment besoin de la notion d'aire d'une partie E du plan (que nous noterons $\mathcal{A}(E)$), dont nous utiliserons de manière essentielle deux propriétés :
1) Additivité.
Si A et B sont deux parties disjointes du plan, on a $\mathcal{A}(A \cup B) = \mathcal{A}(A) + \mathcal{A}(B)$.
2) Invariance par déplacement.
Si u est un déplacement on a $\mathcal{A}(u(E)) = \mathcal{A}(E)$.

Ces deux propriétés peuvent se résumer en disant que l'aire est invariante par découpage et recollement. En fait, nous utiliserons aussi la propriété 1) sous une forme plus forte, dans le cas où les parties A et B ont en commun un ou plusieurs segments (c'est légitime car ceux-ci sont d'aire nulle). De plus, nous verrons au §*g)* qu'on peut se limiter à n'utiliser de la propriété 2) que l'invariance par un seul type de déplacement : la symétrie centrale.

Outre ces propriétés, nous utiliserons quelques notions d'ordre (par exemple l'axiome de Pasch qui dit qu'une droite qui coupe un côté d'un triangle et qui ne passe pas par un sommet en coupe exactement un autre) et la notion de rapports de mesures algébriques sur une droite donnée. En termes modernes, cette notion ne fait que traduire la bijection des droites affines avec $\mathbf{R}$. Nous attirons l'attention du lecteur sur le fait qu'il s'agit là d'une notion affine, puisque toutes les transformations affines conservent ces rapports, même si elles ne conservent pas les distances (considérer par exemple le cas de l'affinité).

a) Les quatre lemmes du collège.

Il s'agit des quatre lemmes suivants, dont nous verrons au paragraphe 3, qu'ils ne font que traduire la semi-invariance de l'aire par certaines transformations du groupe affine.

• Le premier est le lemme "du demi-parallélogramme" qui dit que chaque diagonale partage un parallélogramme en deux triangles de même aire.
C'est une conséquence de l'invariance de l'aire par symétrie centrale.

• Le second est le lemme "du trapèze" qui dit que deux triangles qui ont même base et des sommets sur une parallèle à la base sont d'aires égales.

• Le troisième est le lemme "de la médiane" qui affirme qu'une médiane partage un triangle en deux triangles d'aires égales.

• Le quatrième est le lemme "des proportions" qui dit que si deux triangles ont un sommet commun b et des bases ac et ac' portées par la même droite, le rapport de leurs aires est égal au rapport des bases. Bien entendu le troisième lemme en est un cas particulier.

Ces lemmes, au niveau du collège, peuvent se montrer de manière évidente avec la formule $base \times hauteur\ /2$, [4] ou, mieux encore, par découpage à partir du lemme du demi-parallélogramme, voir §*g)*.

b) Application 1 : Thalès.

Nous allons voir que les quatre lemmes ci-dessus permettent de montrer une quantité de résultats de géométrie affine du triangle. En fait, et c'est ce que nous expliquerons au paragraphe 3, ce qu'affirme la théorie des invariants, c'est que tous les théorèmes affines du triangle peuvent se montrer avec ces lemmes.
Commençons par une preuve du théorème de Thalès :
Soit abc un triangle. Soient $b' \in [ab]$ et $c' \in [ac]$. On supppose $(b'c')$ parallèle à (bc). On a alors $\dfrac{bb'}{ba} = \dfrac{cc'}{ca}$.
En effet, on a $\mathcal{A}(bcc') = \mathcal{A}(bcb')$ par le lemme du trapèze et donc

$$\frac{cc'}{ca} = \frac{\mathcal{A}(bcc')}{\mathcal{A}(abc)} = \frac{\mathcal{A}(bcb')}{\mathcal{A}(abc)} = \frac{bb'}{ba}$$

par le lemme des proportions.

En fait, cette preuve n'est autre que celle d'Euclide et il est piquant (et instructif) que la preuve la plus fidèle, par l'usage qu'elle fait des invariants, à l'esprit d'Erlangen (du moins tel que nous le comprenons et, en tous cas, comme Bourbaki le formule dans le texte cité ci-dessus) soit celle d'Euclide et pas la preuve *via* les vecteurs, ni celles *via* les homothéties ou les projections, seules preuves admises au temps des mathématiques modernes. Entendons-nous bien. Nous n'avons rien contre ces preuves qui sont tout à fait satisfaisantes et qui ont d'autres avantages (notamment de préparer à la prise en compte du fait linéaire). Nous disons simplement que celle proposée ici a, pour le collège, l'avantage de l'économie des moyens puisqu'elle n'utilise que la notion d'aire.

c) Application 2 : médianes.

---

[4] Bien que toutes ces notions soient affines et pas euclidiennes, ce n'est évidemment pas une raison pour se priver des outils euclidiens dans l'enseignement.

Une deuxième application des principes évoqués ci-dessus est la preuve de la concourance des médianes d'un triangle. On montre d'abord qu'un point o (intérieur au triangle) est sur la médiane aa' du triangle abc si et seulement si on a l'égalité des aires $\mathcal{A}(oab) = \mathcal{A}(oac)$. En effet, pour le sens direct on applique deux fois le lemme de la médiane ; pour la réciproque on utilise le lemme suivant, variante immédiate du lemme des proportions, que nous appellerons "lemme du chevron" dans ce qui suit (un dessin explique ce nom si o est un point intérieur du triangle) :

Soit abc un triangle et o un point du plan. Si (oa) coupe (bc) en a', on a la formule

$$\frac{\mathcal{A}(oba)}{\mathcal{A}(oca)} = \frac{a'b}{a'c}.$$

Revenons à la concourance des médianes. Si on note aa', bb', cc' les médianes et g le point d'intersection de (bb') et (cc') on a alors, par le lemme du chevron $\mathcal{A}(gab) = \mathcal{A}(gbc)$ et $\mathcal{A}(gbc) = \mathcal{A}(gac)$, d'où $\mathcal{A}(gab) = \mathcal{A}(gac)$ et le résultat.

d) Application 3 : Ménélaüs.

Rappelons l'énoncé de ce théorème :

Soit abc un triangle et soit D une droite qui coupe respectivement les droites $(bc), (ca), (ab)$ en a', b', c'. On a l'égalité

$$P = \frac{\overline{a'b}}{\overline{a'c}} \times \frac{\overline{b'c}}{\overline{b'a}} \times \frac{\overline{c'a}}{\overline{c'b}} = 1.$$

Il y a deux cas de figures (et deux seulement en vertu de l'axiome de Pasch) : soit deux des points a', b', c', disons b' et c', sont sur les côtés du triangle (figure 1), soit aucun point a', b', c' n'y est (figure 2). Supposons par exemple que les points soient comme sur la figure 1 (le raisonnement est analogue sinon).

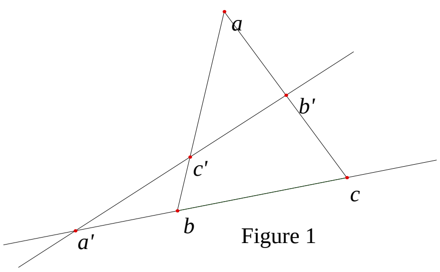


Figure 1

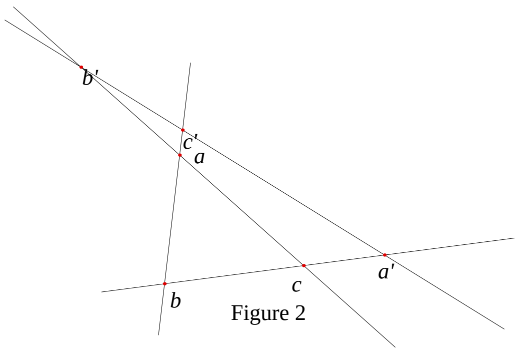


Figure 2

On interprète alors les rapports de la relation comme des rapports d'aires (au signe près) en faisant apparaître deux fois l'aire de $a'bc'$:

$$P = \frac{\mathcal{A}(a'bc')}{\mathcal{A}(a'cc')} \times \frac{\mathcal{A}(b'cc')}{\mathcal{A}(b'ac')} \times \frac{\mathcal{A}(c'aa')}{\mathcal{A}(c'ba')}$$

et on conclut en appliquant les lemmes des proportions et du chevron.

Remarque. En fait, on peut s'épargner la distinction des cas de figures en utilisant l'invariant "aire algébrique", c'est-à-dire, cf. §3, le déterminant des vecteurs ou le produit vectoriel, mais cela suppose de travailler à un niveau plus avancé.

e) Application 4 : Céva et Gergonne.

Les idées évoquées ci-dessus permettent aussi de montrer sans difficulté les théorèmes de Céva et de Gergonne :

Soit abc un triangle et a', b', c' *des points de* $(bc), (ca), (ab)$ *respectivement. Alors, si les droites* (aa'), (bb'), (cc') *concourent en* o *on a les relations :*

$$\frac{\overline{a'b}}{\overline{a'c}} \times \frac{\overline{b'c}}{\overline{b'a}} \times \frac{\overline{c'a}}{\overline{c'b}} = -1 \quad \text{(Céva)} \qquad \text{et} \qquad \frac{\overline{oa'}}{\overline{aa'}} + \frac{\overline{ob'}}{\overline{bb'}} + \frac{\overline{oc'}}{\overline{cc'}} = 1 \quad \text{(Gergonne).}$$

Là encore, il y plusieurs cas de figure à considérer si on travaille avec les aires, qui sont unifiés si on utilise le produit vectoriel. Supposons par exemple que les points a', b', c' sont sur les côtés du triangle. Pour Gergonne on note alors qu'on a $\dfrac{oa'}{aa'} = \dfrac{\mathcal{A}(obc)}{\mathcal{A}(abc)}$ et les relations analogues avec b' et c'. Pour Céva on utilise trois fois le lemme du chevron.

f) D'autres exemples encore.

On peut utiliser encore les quatre lemmes à bien d'autres fins et notamment en géométrie euclidienne puisque l'aire est *a fortiori* un invariant du groupe des isométries. On peut ainsi montrer Pythagore (comme le faisait Euclide) ou encore un certain nombre de relations métriques dans le triangle, en voici un exemple entre mille (avec les notations usuelles) :

$$\frac{a}{\sin \widehat{A}} = \frac{b}{\sin \widehat{B}} = \frac{c}{\sin \widehat{C}}.$$

Le lecteur intéressé pourra consulter à ce sujet la brochure [IP] de l'IREM de Poitiers qui comporte un grand nombre d'exemples intéressants.

Voici encore un petit problème dont la solution est évidente en termes d'aires (ou en utilisant les cas d'isométrie des triangles), mais moins facile si l'on n'emploie que les transformations :

Les hauteurs bb' *et* cc' *d'un triangle abc sont égales. Montrer que abc est isocèle.*

Signalons enfin un thème classique, mais très riche, qui peut faire l'objet de multiples situations de recherche dans des classes et où les quatre lemmes seront très efficaces. Il s'agit du partage de polygones en parties de même aire ou ayant des rapports d'aires donnés, cf. [F]. En voici un exemple simple :

On considère un triangle T *et un point* m *situé sur un côté de* T*. Comment partager* T *en deux parties de même aire par une droite issue de* m *?*

On peut ensuite demander de partager le triangle en trois, ou passer au cas d'un quadrilatère, etc.

g) Un mot sur le découpage.

Nous avons dit plus haut que les lemmes "du collège" peuvent tous être prouvés en utilisant la formule *base* × *hauteur* / 2.

Il y a toutefois de bonnes raisons didactiques, s'agissant de l'introduction des aires à l'école élémentaire et au collège, de ne pas écraser trop vite la notion d'aire sur

les nombres, cf. [DPG]. Cela peut conduire à préférer les procédures de découpage et recollement à un usage trop exclusif des formules.
De plus, nous allons voir que cette procédure de découpage est très efficace puisqu'elle permet de montrer tous les lemmes évoqués ci-dessus à partir du seul lemme du demi-parallélogramme (i.e. l'invariance par symétrie centrale).
A posteriori, ce phénomène a une justification théorique intéressante qui vient du théorème de Bolyai, revu par Hadwiger-Glur (cf. annexe 2 ou [B1] et [B2]). En effet, ce théorème affirme que deux polygones de même aire sont toujours équivalents par découpage et recollement (c'est-à-dire par puzzle) et qu'il suffit même d'utiliser, pour déplacer les pièces, des translations et des symétries centrales. Comme toute translation est elle-même composée de deux symétries centrales on peut même utiliser seulement les symétries. Si P et Q sont deux polygones de même aire, on peut donc passer du polygone P au polygone Q en découpant P en un nombre fini de polygones P_i, puis en déplaçant les P_i par des symétries centrales et en les utilisant pour reconstituer Q.

Si l'on croit en ce résultat, on doit pouvoir montrer le lemme du trapèze et celui de la médiane par découpage à partir du lemme du demi-parallélogramme. C'est effectivement le cas (et c'est à peu de choses près ainsi que procède Euclide !). On montre en effet, successivement (le lecteur se reportera aux figures qui parlent souvent d'elles-mêmes) :

• qu'un triangle (ici mcd, cf. figure 3) dont la base est l'un des côtés d'un parallélogramme et dont le sommet est sur le côté opposé a pour aire la moitié de celle du parallélogramme : on trace la parallèle à (ad) par m qui recoupe (cd) en n et on applique deux fois le lemme du demi-parallélogramme,

• le lemme du trapèze, cf. figure 4 (il y a un autre cas de figure plus facile) : on a $\mathcal{A}(abc) = \frac{1}{2}\,\mathcal{A}(aa'b'b) - \mathcal{A}(a'cb')$ par la remarque précédente et cette aire est encore égale à $\mathcal{A}(a'bc)$ par le lemme du demi-parallélogramme,

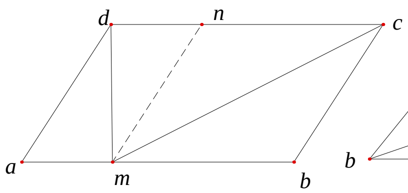


Figure 3

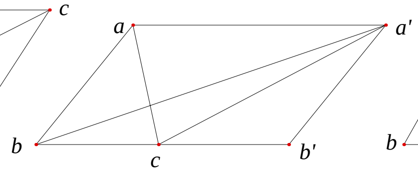


Figure 4

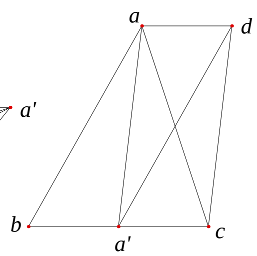


Figure 5

• le lemme de la médiane, cf. figure 5 : les deux parallélogrammes $aba'd$ et $aa'cd$ ont même aire (ils ont une moitié commune : ada'), donc leurs moitiés aba' et aca' ont aussi même aire.
Il reste à montrer le lemme des proportions. Là on n'évite pas un problème de passage à la limite. On commence par montrer le lemme lorsque le rapport $r = \dfrac{ac}{a'c}$ vaut $\dfrac{1}{2^n}$ par application répétée du lemme de la médiane. On en déduit que le lemme des proportions vaut lorsque r est de la forme $\dfrac{k}{2^n}$ par additivité de l'aire. Pour conclure il reste à utiliser la croissance de l'aire (qui résulte de l'additivité) pour obtenir un encadrement et passer à la limite. Bien entendu ce phénomène

est inévitable et il est implicitement présent déjà dans un calcul aussi familier que celui de l'aire du rectangle comme produit de la longueur par la largeur et donc aussi dans le calcul de l'aire du triangle.

Entendons-nous bien : même si nous considérons que les méthodes de découpage-recollement sont satisfaisantes car elles permettent d'obtenir tous les résultats visés avec une remarquable économie de moyens, nous ne disons pas qu'il est nécessaire de faire toutes ces démonstrations au collège. En effet, la formule *base* × *hauteur* / 2 a le mérite de permettre des conclusions rapides et convaincantes. En revanche, s'il s'agit de construire la notion d'aire et d'apprendre à raisonner, initier les élèves à utiliser ce type de preuves dynamiques nous semble un objectif intéressant.

3. Explication du succès de la méthode : la théorie des invariants.

Il s'agit de la théorie des invariants polynomiaux associés à un groupe. Cette théorie a été fondée vers le milieu du dix-neuvième siècle par Cayley, Sylvester et Hermite. Dans ce paragraphe nous allons en expliquer le principe dans le cas de la géométrie affine plane afin de montrer que l'efficacité de l'invariant aire constatée au paragraphe précédent, n'est pas due au hasard, mais qu'elle est totalement justifiée par la théorie.
Le lecteur est averti que le paragraphe qui suit, et qui est de nature algébrique, va être un peu plus technique (encore que les démonstrations délicates soient rejetées dans l'annexe 1). Cependant, ces considérations sont indispensables pour comprendre les raisons profondes de l'importance de l'aire en géométrie affine.

Dans ce paragraphe nous travaillerons dans un plan affine réel, associé à un espace vectoriel de dimension 2 et muni d'un repère affine : un point o et deux vecteurs indépendants e_1, e_2. On peut alors identifier un point x du plan avec le vecteur $\overrightarrow{ox}$ ou encore avec le couple de ses coordonnées : $x = (x_1, x_2)$. Nous ferons systématiquement ces identifications dans ce qui suit.

a) Action sur les polynômes.

Lorsqu'on travaille en géométrie affine plane on considère au départ des points donnés du plan. Prenons l'exemple le plus simple : celui de la géométrie d'un triangle a, b, c. Ces points sont décrits par leurs coordonnées dans le repère o, e_1, e_2 : $a = (a_1, a_2)$, $b = (b_1, b_2)$, $c = (c_1, c_2)$ et la plupart des constructions "affines" usuelles de points du plan à partir des données (notamment par intersections de droites, voir ci-dessous) conduisent à des points dont les coordonnées sont des fonctions polynomiales ou rationnelles en les a_i, b_i, c_i. (5)

Le groupe affine du plan (engendré par les translations et les applications linéaires bijectives du plan dans lui-même) agit alors sur les points donnés. Quitte à faire

(5) On supposera en général les points a, b, c "généraux". Cela signifie qu'ils ne sont pas liés par des relations algébriques, par exemple qu'ils ne sont pas alignés. En vertu du théorème de prolongement des identités algébriques on peut alors identifier fonctions polynomiales et polynômes en a, b, c et donc considérer les a_i, b_i, c_i comme des indéterminées.

une translation, on se ramène au cas où l'origine est fixe,([6]) de sorte que le groupe qui agit est alors $GL(2) = GL(2, \mathbf{R})$, groupe des applications linéaires inversibles (ou des matrices 2×2 inversibles) avec lequel nous travaillerons désormais. Ce groupe agit simplement sur les coordonnées des points a, b, c par l'intermédiaire des matrices, ainsi, si u a pour matrice $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, on a

$$u(a_1, a_2) = (\alpha a_1 + \beta a_2, \gamma a_1 + \delta a_2).$$

On en déduit une action de $GL(2)$ sur l'anneau de polynômes $\mathbf{R}[a_1, a_2]$ (et sur le corps des fractions rationnelles, ou plus généralement sur les fonctions de a_1 et a_2) définie au moyen des formules ci-dessus. Précisément, si P est un polynôme et u un élément de $GL(2)$, u agit sur P ([7]) par :

$$(P.u)(a) = (P \circ u)(a) = P(u(a)), \quad \text{c'est-à-dire} \quad (P.u)(a_1, a_2) = P(u(a_1, a_2)),$$

ou encore, si u admet la matrice ci-dessus, $(P.u)(a_1, a_2) = P(\alpha a_1 + \beta a_2, \gamma a_1 + \delta a_2)$. Bien entendu $GL(2)$ opère aussi sur $\mathbf{R}[a_1, a_2, b_1, b_2, c_1, c_2]$ en posant

$$(P.u)(a, b, c) = (P.u)(a_1, a_2, b_1, b_2, c_1, c_2) = P\big(u(a_1, a_2), u(b_1, b_2), u(c_1, c_2)\big)$$

et on a évidemment une définition analogue avec un nombre quelconque de points.

Donnons quelques exemples concrets de cette opération en prenant pour P le polynôme $a_1^2 + a_2^2$ qui, dans le cas euclidien, correspond au carré de la norme du vecteur a. Si u est l'affinité de matrice $\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}$, on a $P.u = a_1^2 + \lambda^2 a_2^2$; si v est l'application de matrice $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$ (qu'on appelle une transvection), on a $P.v = a_1^2 + 2\lambda a_1 a_2 + (\lambda^2+1) a_2^2$; si w est l'application de matrice $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ (une rotation) un petit calcul donne $P.w = P$.

b) Invariants.

On désigne simplement par le mot **invariants** les polynômes (ou les fractions rationnelles, voire les fonctions) invariants par cette action (ou par celle d'un sous-groupe). On voit, par exemple, que le polynôme $a_1^2 + a_2^2$ est invariant par rotation, mais ni par affinité, ni par transvection.

([6]) En fait, les notions auxquelles nous nous intéressons étant toutes de nature vectorielle (elles dépendent des vecteurs et non des points, penser à la longueur, l'angle ou l'aire), elles sont automatiquement conservées par les translations qui ne jouent donc aucun rôle dans la théorie.

([7]) Telle que nous venons de la définir, il s'agit d'une opération à droite. Si l'on souhaite avoir une opération à gauche, ce qui est plus habituel, il faut poser

$$(u.P)(a_1, a_2; b_1, b_2; c_1, c_2) = P\big(u^{-1}(a_1, a_2); u^{-1}(b_1, b_2); u^{-1}(c_1, c_2)\big).$$

Nous avons choisi ici l'opération à droite pour éviter d'avoir trop d'exposants -1. On notera que cela ne change pas les invariants.

On trouve, parmi les invariants du groupe $SL(2)$ des applications linéaires de déterminant 1, les déterminants $a \wedge b = a_1 b_2 - a_2 b_1$ et de même $b \wedge c$ et $c \wedge a$. En effet, si u a pour matrice $U = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, de déterminant 1 et si A est la matrice $\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$, on vérifie qu'on a $a \wedge b = \text{dét}\, A$ et $(a \wedge b).u = \text{dét}\,(UA) = \text{dét}\, A = a \wedge b$ (on peut aussi faire la vérification à la main). Sous l'action du groupe $GL(2)$ tout entier on trouve notamment comme invariants les quotients du type $\dfrac{a \wedge b}{b \wedge c}$. En effet, si on applique une transformation u de déterminant λ, le numérateur et le dénominateur sont multipliés par λ, de sorte que la fraction est invariante.

Ces invariants ont, bien entendu, un sens géométrique. De fait, la quantité $\frac{1}{2}(a \wedge b)$ n'est autre que l'aire algébrique du triangle oab, comptée positivement si la base (a, b) est directe, négativement sinon. Il n'est pas difficile de se convaincre de ce fait en utilisant le découpage et recollement. On peut aussi le comprendre, dans le cadre euclidien, en interprétant le déterminant $a \wedge b$ comme produit vectoriel des vecteurs a et b.

c) Retour aux lemmes du collège.

Nous venons de montrer algébriquement que l'aire est un semi-invariant dans l'action du groupe affine. En réalité, c'est ce que nous avions établi géométriquement au paragraphe précédent avec les quatre lemmes du collège. ([8])
En effet, prenons dans le plan le repère affine a, b, c (avec a comme origine). Parmi les transformations de $GL(2)$ il y a en particulier les suivantes :
Il y a d'abord les transvections (certains auteurs parlent de cisaillements, traduction de l'anglais *shear*), qui correspondent, à changement de base près, aux matrices $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$. Géométriquement, une telle transformation fixe a et b et déplace c sur une parallèle à (ab).
La traduction de l'invariance de l'aire par transvection n'est autre que le lemme du trapèze.

Il y a ensuite les dilatations (ou affinités), de matrices $\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}$. Géométriquement, une telle transformation fixe encore a et b, mais déplace c sur la droite (ac). La traduction de la semi-invariance de l'aire par dilatation n'est autre que le lemme des proportions.
Un cas particulier de dilatation est la symétrie oblique d'axe (ab) et de direction (ac) qui correspond à $\lambda = -1$. La traduction de l'invariance de l'aire par symétrie oblique n'est autre que le lemme de la médiane.
Enfin, pour faire bonne mesure, on peut aussi ajouter à ces éléments la symétrie centrale, de matrice $-\text{Id}$, pour laquelle la traduction de l'invariance est le lemme du demi-parallélogramme.
Comme le groupe $GL(2)$ est engendré par les dilatations et les transvections (voir par exemple [P] Chapitre 4, ou faire un petit calcul à la main) on voit que les

([8]) En fait, ce que nous avons prouvé c'est l'invariance de l'aire ordinaire, valeur absolue de l'aire orientée, et c'est ce qui nous a amené à distinguer des cas de figures pour exprimer les aires tantôt comme sommes, tantôt comme différences.

lemmes du collège montrent en fait toute la semi-invariance. ([9])

d) La théorie : premier point.

Le premier point fondamental de la théorie des invariants, dans le cas du groupe linéaire comme d'ailleurs de tous les groupes dits "classiques" qui gouvernent les géométries usuelles, c'est qu'on sait déterminer explicitement **tous** les invariants. Dans le cas ci-dessus, on démontre (mais ce n'est pas un résultat tout à fait trivial, cf. annexe 1 ou [W]) que les invariants du groupe $SL(2, \mathbf{R})$ s'expriment polynomialement à partir des trois produits $a \wedge b$, $b \wedge c$ et $c \wedge a$. Pour le groupe $GL(2, \mathbf{R})$ il n'y a pas d'invariants polynomiaux non constants, mais les fractions rationnelles invariantes sont des quotients de deux polynômes homogènes de même degré invariants sous $SL(2)$.

Ce résultat est la confirmation de ce que nous avions avancé au paragraphe 1 : l'invariant aire (ou rapport d'aire) est bien, dans le cas de la géométrie affine d'un triangle, le seul possible. ([10])
Comme les quatre lemmes du collège montrent cette invariance, ils épuisent donc la théorie sur ce point.

e) Relations et théorèmes.

Nous en venons au second point de la théorie : le lien entre les invariants et les théorèmes. Reprenons le cas de la géométrie affine, avec une origine o et trois points a, b et c.

Idée 1 : points et invariants.

Nous avons d'abord besoin de préciser un peu formellement ce qu'on peut appeler une construction affine.

D'abord, effectuer une construction c'est produire, à partir des points de départ a, b, c un nouveau point m. Autrement dit, une construction est une application F qui associe aux points de départ a, b, c un certain point $m = F(a, b, c)$. Ensuite, dire qu'on a affaire à une construction affine signifie que cette application est résistante aux changements de repères affines, c'est-à-dire aux bijections affines. Si g est une application affine bijective du plan (ou même linéaire si on continue à fixer l'origine o conformément à nos principes), la construction F doit donner la

([9]) On peut même oublier les transvections car elles sont produit de deux symétries :

$$\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & \lambda \\ 0 & -1 \end{pmatrix}.$$

Le lecteur avisé pourra donc, à titre d'exercice, montrer qu'on retrouve le lemme du trapèze à partir de celui de la médiane.

([10]) Le résultat est d'ailleurs valable pour un nombre quelconque de points. Dans la théorie on considère aussi, à côté des vecteurs, des directions de droites, c'est-à-dire des formes linéaires sur le plan vectoriel. On a alors de nouveaux invariants de deux types : soit $f(a)$, si f est une forme et a un vecteur, soit le déterminant $f \wedge g$, si f et g sont deux formes. Pour certains exemples géométriques, par exemple celui de Ménélaüs, les démonstrations sont plus commodes en utilisant à la fois vecteurs et formes, mais cela déborde du cadre de notre propos.

même chose après transport par g. En termes mathématiques cela signifie qu'on peut parcourir le diagramme suivant dans n'importe quel sens :

$$\begin{array}{ccc} a,b,c & \xrightarrow{F} & F(a,b,c) \\ \Big\downarrow g & & \Big\downarrow g \\ g(a),g(b),g(c) & \xrightarrow{F} & F(g(a),g(b),g(c)) = g(F(a,b,c)), \end{array}$$

c'est-à-dire que F et g commutent : $F(g(a),g(b),g(c)) = g(F(a,b,c))$.
Le lecteur vérifiera que c'est le cas pour toutes les constructions usuelles de la géométrie affine,([11]). Par exemple, si F consiste à prendre le milieu m de $[a,b]$, et si on transforme a,b par g, $g(m)$ sera bien le milieu de $[g(a),g(b)]$.

Par ailleurs, comme les points b et c sont généraux, l'origine o et les points b,c ne sont pas alignés, donc forment un repère affine du plan. On peut donc écrire, **dans ce repère associé au triangle** $F(a,b,c) = f_1(a,b,c)\,b + f_2(a,b,c)\,c$ avec $f_i(a,b,c) \in \mathbf{R}$. On a alors le résultat suivant, évident mais fondamental (et qui vaut dans un cadre beaucoup plus général) :

La construction $m = F(a,b,c)$ est affine si et seulement si les fonctions f_1 et f_2 sont invariantes sous l'action de $GL(2,\mathbf{R})$.

Démonstration. Soit $g \in GL(2,\mathbf{R})$. Comme g est linéaire, on a $g(m) = f_1(a,b,c)\,g(b) + f_2(a,b,c)\,g(c)$. Or, l'hypothèse que la construction est affine donne la relation :

$$g(m) = f_1(g(a),g(b),g(c))\; g(b) + f_2(g(a),g(b),g(c))\; g(c),$$

mais, comme $g(b),g(c)$ est une base du plan vectoriel, on en déduit $f_1(a,b,c) = f_1(g(a),g(b),g(c))$ et $f_2(a,b,c) = f_2(g(a),g(b),g(c))$, cqfd.

Autrement dit, les coordonnées dans le repère o,b,c des points construits, par des constructions affines, à partir des points a,b,c sont des polynômes ou des fractions rationnelles en les indéterminées a_i,b_i,c_i (et en tous cas des fonctions des a_i,b_i,c_i) qui sont des **invariants**. Donnons quelques exemples, le lecteur pourra en produire de nombreux autres. Le point a lui-même s'écrit dans ce repère

$$a = \frac{c\wedge a}{c\wedge b}\; b + \frac{a\wedge b}{c\wedge b}\; c\,,$$

(il suffit d'écrire $a = \lambda b + \mu c$ et d'appliquer $.\wedge b$ et $.\wedge c$ pour calculer λ et μ) tandis que le point b', intersection de (ob) et (ac) s'écrit

$$b' = \frac{-c\wedge a}{a\wedge b + b\wedge c}\; b$$

et le point c'', intersection de (ab) et de la parallèle à (bc) passant par b' s'écrit :

$$c'' = \frac{a\wedge b + a\wedge c}{a\wedge b + b\wedge c}\, b - \frac{a\wedge b}{a\wedge b + b\wedge c}\, c.$$

Idée 2 : théorèmes et relations.

([11]) On entend notamment par là les opérations suivantes : construction de la droite passant par deux points, intersection de deux droites, construction de parallèles, de milieux et de barycentres, etc.

On montre facilement que toutes les propriétés usuelles de géométrie affine (alignement de points, concourance de droites, rapports de mesures algébriques) s'expriment comme des **relations**, polynomiales elles aussi, entre les coordonnées des points en question dans un repère quelconque (ou entre les coefficients des équations des droites). Par exemple, dire que les points $x = (x_1, x_2)$, $y = (y_1, y_2)$, $z = (z_1, z_2)$ sont alignés c'est dire que le déterminant D_1 ci-dessous est nul, tandis que dire que les trois droites $u_i x + v_i y + w_i = 0$, $(i = 1, 2, 3)$ sont concourantes c'est dire que le déterminant D_2 est nul :

$$D_1 = \begin{vmatrix} 1 & 1 & 1 \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}, \qquad D_2 = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix}.$$

Dans les deux cas on a bien des relations polynomiales entre les coordonnées ou les coefficients.
Comme on a vu que les coordonnées des points (dans le repère o, b, c) sont des invariants on peut donc formuler le principe suivant :

Principe : *Les théorèmes de la géométrie affine du triangle s'obtiennent comme relations polynomiales entre les invariants du groupe* $GL(2, \mathbf{R})$.

Exemple : les médianes.

Pour illustrer ce principe, voici la démonstration en termes de relations entre invariants, de la concourance des médianes du triangle.

On considère un triangle a, b, c. Dire que l'origine o du plan est sur la médiane issue de a signifie qu'on a $a \wedge (b + c) = 0$. ([12]) Or, on a la relation (triviale) entre les invariants

$$a \wedge (b + c) + b \wedge (c + a) + c \wedge (a + b) = 0$$

qui traduit la bilinéarité et l'antisymétrie du déterminant, de sorte que si on prend pour origine l'intersection de deux des médianes, elle est aussi sur la troisième.

f) La théorie : deuxième point.

L'intérêt d'avoir identifié les théorèmes d'une géométrie et les relations entre les invariants de cette géométrie vient alors du second résultat fondamental de la théorie des invariants. En effet, ce résultat affirme que, dans toutes les géométries usuelles, non seulement les invariants sont connus, mais encore les relations éventuelles entre eux le sont aussi. Dans le cas de la géométrie affine de trois points, on démontre (ce n'est pas très difficile, cf. annexe 1) qu'il n'y a pas de relations non triviales entre les trois invariants $a \wedge b, b \wedge c, c \wedge a$. En revanche, à partir de 4 points a, b, c, d, si on a toujours les mêmes invariants, il apparaît des relations non triviales du type :

$$(b \wedge c)(a \wedge d) + (c \wedge a)(b \wedge d) + (a \wedge b)(c \wedge d) = 0.$$

([12]) Si on préfère écrire avec des vecteurs munis de flèches, cette égalité devient $\overrightarrow{oa} \wedge (\overrightarrow{ob} + \overrightarrow{oc}) = \vec{0}$. Si on pose $\overrightarrow{om} = \overrightarrow{ob} + \overrightarrow{oc}$, la relation signifie que o, a, m sont alignés. Or, $obmc$ est un parallélogramme, de sorte que (om) passe par le milieu a' de $[bc]$ et on a bien le résultat. Cette relation n'est autre que la version algébrique de l'égalité d'aires $\mathcal{A}(oab) = \mathcal{A}(oac)$ utilisée au paragraphe 2.

En effet, cette relation s'obtient à partir de la formule $(b\wedge c)a+(c\wedge a)b+(a\wedge b)c=0$ que nous avons vue ci-dessus en appliquant $.\wedge d$. À notre connaissance, cette relation n'a pas de traduction très spectaculaire dans le cadre de la géométrie affine du plan réel. En revanche, elle se traduit en termes de birapports par la relation bien connue $[abcd]+[acbd]=1$. Si on travaille sur les complexes, on en déduit en particulier le théorème de Ptolémée : si un quadrilatère convexe *abcd* est inscrit dans un cercle le produit de ses diagonales est égal à la somme des produits de ses côtés opposés.

g) Bilan.

Ainsi, ce qu'affirment le programme d'Erlangen et la théorie des invariants, c'est qu'on peut, en principe, obtenir tous les théorèmes de géométrie à partir de la connaissance des invariants et de leurs relations éventuelles. Or, dans le cas de la géométrie affine du triangle, nous avons vu qu'il n'y a qu'un invariant : l'aire et aucune relation. C'est ce qui explique que les quatre lemmes du collège qui expriment que l'aire est un (semi)-invariant sous le groupe affine permettent de prouver tous les théorèmes du domaine considéré.

4. Conclusion.

L'unique conclusion que nous souhaitons retenir de cette étude sur le programme d'Erlangen et les invariants c'est l'importance capitale de ceux-ci. Il est assez clair que la plupart des promoteurs de la réforme des mathématiques modernes avaient en tête le programme d'Erlangen, mais pas le codicille que constitue la théorie des invariants, ce qui les a conduit à occulter ceux-ci au profit des seules transformations et cet état de fait perdure encore, sous une forme atténuée certes, de nos jours. C'est un point sur lequel le rapport de la commission Kahane insiste : il y a un grand intérêt, notamment pour l'enseignement de la géométrie au collège, à utiliser les invariants et pas seulement les transformations. Nous espèrons que les quelques exemples développés ici le montrent suffisamment. De plus en plus de professeurs en sont convaincus par leur pratique, mais, formés à l'usage quasi-exclusif des transformations, ils ne se sentent pas encore autorisés à prendre un peu de recul par rapport à celles-ci et sont souvent réticents envers des preuves utilisant les aires comme celles vues ci-dessus, ou les angles en géométrie euclidienne. On constate d'ailleurs le même phénomène par rapport aux cas d'"égalité" des triangles. Nous voudrions ici les rassurer en leur apportant une sorte de caution théorique.

Un autre point est apparu clairement dans ce qui précède : c'est l'importance de la figure. En effet, les démonstrations géométriques utilisant les aires nécessitent souvent de distinguer plusieurs cas de figure. C'est la rançon du fait de n'utiliser que la valeur absolue de l'invariant aire algébrique, c'est-à-dire du déterminant ou du produit vectoriel. En contrepartie, l'avantage de l'utilisation des aires c'est son côté visuel, qui manque cruellement quand on manipule produits vectoriels et relations de Chasles de manière purement algébrique. On a alors le sentiment d'une activité purement mécanique, comme le notait déjà Jean-Jacques Rousseau : *Je n'ai jamais été assez loin pour bien sentir l'application de l'algèbre à la géométrie. Je n'aimais pas cette manière d'opérer sans voir ce qu'on fait, et il me semblait que résoudre un problème de géométrie par les équations, c'était jouer un air en tournant une manivelle.*

Nous considérons pour notre part ce passage par les figures, avec les difficultés inhérentes aux divers cas à envisager, comme une étape inévitable et nécessaire au collège. L'unification qu'apportent les calculs vectoriels, pour intéressante qu'elle soit, doit venir à son heure, c'est-à-dire, sans doute, au lycée. Ensuite, par rapport aux difficultés liées notamment à la convexité, qui est souvent sous-jacente dans ce type de problèmes, notre opinion est que les professeurs doivent maîtriser ces questions, mais qu'il est inutile, voire néfaste, de soulever trop tôt ces difficultés pour les jeunes élèves. L'objectif de la géométrie au collège n'est certes pas de s'échiner à montrer, avec peine, des choses que l'on voit aussitôt sur la figure, mais plutôt d'utiliser des outils intuitifs et efficaces comme les notions de longueur, d'angle et d'aire et un premier apprentissage du raisonnement, pour établir des théorèmes non évidents.

Pour terminer nous voudrions ajouter un point important. Nous espérons évidemment avoir convaincu le lecteur de l'intérêt et du bien fondé de l'utilisation de la notion d'aire au collège comme outil de démonstration. Toutefois nous serions désolés s'il en déduisait qu'hors des aires il n'est point de salut et courait jeter aux orties toutes les autres démonstrations de Thalès et des autres. Plus que toute méthode, si pertinente soit-elle, la variété des approches ne fait-elle pas tout le charme de notre discipline ?

Références

[B1] Boltianskii V., Equivalent and equidecomposable figures, D.C. Heath and company, 1963.

[B2] Boltianskii V. Hilbert's third problem, Winston & sons, 1978.

[Bo] Bourbaki N. Éléments d'histoire des mathématiques, Hermann, 1960.

[D] Dieudonné J., Algèbre linéaire et géométrie élémentaire, Hermann, 1968.

[DPG] Douady R. et Perrin-Glorian M.-J., Aires de surfaces planes, *Petit x* numéros 6, p. 5-33 et numéro 8 p. 5-30, IREM de Grenoble, 1984-85.

[F] Fourrey E., Curiosités géométriques, Vuibert, Paris 1907, réédité en 1994.

[G] Guinot M., Le paradoxe de Banach-Tarski, Aleas, Lyon, 1991.

[IP] Aires second cycle, Brochure de l'IREM de Poitiers, 1991.

[K] Klein F., Le programme d'Erlangen, coll. Le discours de la méthode, Gauthier-Villars, 1969.

[L] Lebesgue H., La mesure des grandeurs, librairie Blanchard, 1975.

[P] Perrin D., Cours d'algèbre, Ellipses, 1996.

[W] Weyl H., Invariant theory, Princeton mathematical Series, Princeton, 1946.

Remerciements.

Il m'est agréable de remercier ici plusieurs personnes qui ont permis que cet article voie le jour. Je remercie d'abord Henri Bareil, qui est à l'origine de cette publication, et Jean-Pierre Kahane, pour m'avoir autorisé à publier un texte qui s'appuie largement sur des réflexions que nous avons

menées dans la commission qu'il préside. Je remercie vivement Jean-Claude Duperret et Jean-Pierre Richeton pour le travail que nous avons effectué ensemble dans cette commission, pour leur lecture attentive de cet article et pour leur souci permanent de toucher le plus grand nombre de professeurs. Je remercie aussi Marie-Claude David et Pierre Legrand pour leurs remarques. Enfin, ce texte est en réalité le fruit d'innombrables discussions avec mon épouse Marie-Jeanne Perrin-Glorian qui s'est penchée bien avant moi sur cette notion d'aire et a su me convaincre de son intérêt. Le lecteur doit à l'acuité de sa vision géométrique une bonne partie des lemmes ci-dessus et de leurs démonstrations. Pour tout cela et pour bien d'autres choses encore, je la remercie chaleureusement.

Annexe 1 : Les invariants du groupe $GL(2)$

0. Introduction.

Dans cette annexe on détermine les invariants des groupes $GL(2, \mathbf{R})$ et $SL(2, \mathbf{R})$ dans leur action sur les polynômes en six variables $a_1, a_2; b_1, b_2; c_1, c_2$ correspondant à trois vecteurs a, b, c de $\mathbf{R}^2$. Ce cadre est essentiellement celui de la géométrie affine du triangle évoquée dans le corps du texte. Le lecteur qui désirerait en savoir plus sur le cas général pourra consulter [W] (mais cet ouvrage, déjà ancien, n'est pas d'une lecture facile).

1. Les opérations et les invariants.

a) Opérations.

L'anneau de polynômes $A = \mathbf{R}[a_1, a_2; b_1, b_2; c_1, c_2]$ sera noté aussi $\mathbf{R}[a, b, c]$ en posant $a = (a_1, a_2)$, $b = (b_1, b_2)$ et $c = (c_1, c_2)$. On considère le groupe linéaire $GL(2, \mathbf{R}) = GL(2)$ des matrices 2×2 à coefficients réels. On rappelle que si u est la matrice $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, on a la formule $u(x_1, x_2) = (\alpha x_1 + \beta x_2, \gamma x_1 + \delta x_2)$.

Le groupe $GL(2)$ admet une opération naturelle sur l'anneau A obtenue en posant, pour $u \in GL(2, \mathbf{R})$ et $P \in A$:

$$(P.u)(a_1, a_2; b_1, b_2; c_1, c_2) = P\big(u(a_1, a_2); u(b_1, b_2); u(c_1, c_2)\big).$$

On peut aussi considérer cette opération en restriction à un sous-groupe G de $GL(2)$, notamment au sous-groupe $SL(2)$ des matrices de déterminant 1.

Les opérations ci-dessus s'étendent immédiatement au corps des fractions rationnelles $\mathbf{R}(a, b, c)$ en posant $\dfrac{P}{Q}.u = \dfrac{P.u}{Q.u}$.

b) Invariants.

Définition 1.1.
Soit G un sous-groupe de $GL(2)$ (par exemple $G = GL(2)$ ou $G = SL(2)$).
1) On appelle **invariant** *sous G un polynôme (ou une fraction rationnelle) P qui vérifie $P.u = P$ pour tout $u \in G$.*
2) On appelle **invariant relatif** *sous G un polynôme (ou une fraction rationnelle) P qui vérifie : pour tout $u \in G$ il existe $\lambda(u) \in \mathbf{R}^*$ tel que $P.u = \lambda(u)P$.*

Remarques 1.2.
a) Il est clair que l'ensemble des invariants est une sous-algèbre de $\mathbf{R}[a, b, c]$.
b) Le cas des invariants relatifs sera étudié au §4. On verra que le multiplicateur $\lambda(u)$ est lié au déterminant de u.

Proposition 1.3.
Les polynômes $b \wedge c = b_1c_2 - b_2c_1$, $c \wedge a = c_1a_2 - c_2a_1$ et $a \wedge b = a_1b_2 - a_2b_1$ sont des invariants sous $SL(2)$ et des invariants relatifs sous $GL(2)$.

Démonstration. Soit $m(b, c)$ la matrice $\begin{pmatrix} b_1 & c_1 \\ b_2 & c_2 \end{pmatrix}$. Si u est une matrice de $GL(2)$ on a $m(u(b), u(c)) = u\ m(b, c)$. Comme $b \wedge c$ n'est autre que $\text{dét}\, m(b, c)$ on en déduit $(b \wedge c).u = (\text{dét}\, u)(b \wedge c)$, ce qui prouve l'assertion.

c) Quelques propriétés des invariants.

Désignons par $x = (x_1, x_2)$ l'un des couples d'indéterminées a, b, c. Un polynôme $P(a, b, c)$ sera dit homogène de degré d en x s'il est de la forme $\sum_{i_1+i_2=d} \alpha_{i_1,i_2} x_1^{i_1} x_2^{i_2}$ où les α_{i_1,i_2} sont des polynômes en les variables autres que les x_i. On a alors le résultat suivant:

Proposition 1.4.
Soit $P \in \mathbf{R}[a, b, c]$ un polynôme.
1) Le polynôme P s'écrit de manière unique sous la forme $P = \sum_{i,j,k} P_{i,j,k}$ où la somme est étendue aux indices $0 \leq i \leq p$, $0 \leq j \leq q$, $0 \leq k \leq r$ et où $P_{i,j,k}$ est homogène de degré i en a_1, a_2, j en b_1, b_2 et k en c_1, c_2.
2) Le polynôme P est invariant sous $SL(2)$ ou $GL(2)$ si et seulement si tous les $P_{i,j,k}$ le sont.

Démonstration. Pour l'existence de l'écriture il suffit de regrouper dans P tous les monômes des degrés considérés.
L'assertion sur les invariants résulte du fait que l'opération conserve les degrés partiels en a, b, c.

2. Polarisation.

a) Définition.

Il s'agit d'une opération sur les polynômes, liée à la dérivation, qui va jouer un rôle crucial dans le calcul des invariants. Cette notion a un sens général, mais nous contentons de la définir dans le cadre restrictif dans lequel nous nous sommes placés.

Définition 2.1.
Soit $f \in \mathbf{R}[a, b, c]$. On désigne par x, y deux des lettres a, b, c (éventuellement égales), de sorte que x et y sont des couples (x_1, x_2), (y_1, y_2). On définit le polynôme $D_{yx}f \in \mathbf{R}[a, b, c]$ par la formule :

$$D_{yx}f(a, b, c) = y_1 \frac{\partial f}{\partial x_1}(a, b, c) + y_2 \frac{\partial f}{\partial x_2}(a, b, c).$$

Remarques 2.2.
a) Les applications D_{yx} sont des endomorphismes de l'espace vectoriel $\mathbf{R}[a,b,c]$ appelés polarisations.
b) Si f ne dépend pas de x on a $D_{yx}f = 0$. Si f est un polynôme homogène de degré p en les variables x et q en les variables y l'opération D_{yx} transforme f en un polynôme homogène de degré $p-1$ en x et $q+1$ en y.
c) Si f est un polynôme homogène de degré p en x, on a $D_{xx}f = pf$.
d) Soit $f \in \mathbf{R}[a,b,c]$. Considérons f comme un polynôme en x. Le polynôme $D_{yx}f$ est le coefficient du premier degré dans le développement de Taylor de $f(x+ty)$, pour $t \in \mathbf{R}$:

$$f(x+ty) = f(x) + tD_{yx}f(x,y) + t^2 \cdots$$

b) Polarisation et invariants.

L'intérêt de l'opération de polarisation est qu'elle transforme invariants en invariants. Dans ce qui suit G désigne un sous-groupe de $GL(2)$.

Proposition 2.3.
Soit $f \in \mathbf{R}[a,b,c]$ un invariant sous G. Alors $D_{yx}(f)$ est invariant sous G.

Démonstration. Soit $u \in G$. En vertu de l'invariance de f on a, pour $t \in \mathbf{R}$:

$$f(u(a+tb), u(b), u(c)) = f(a+tb, b, c),$$

soit encore, puisque u est linéaire, $f(u(a)+tu(b), u(b), u(c)) = f(a+tb, b, c)$. On développe ces deux termes par la formule de Taylor et on obtient, au premier ordre :

$$f(u(a), u(b), u(c)) + tD_{ab}f(u(a), u(b), u(c)) = f(a,b,c) + tD_{ab}f(a,b,c).$$

Comme f est invariant, on voit que $D_{ab}f$ l'est aussi.
Précisons ce que deviennent les invariants de la proposition 1.3 :

Proposition 2.4.
Les polynômes $b \wedge c$, $c \wedge a$, $a \wedge b$ sont permutés (au signe près) ou transformés en 0 par polarisation. Précisément, on a les formules : $D_{ab}(a \wedge b) = D_{ba}(a \wedge b) = 0$, $D_{aa}(a \wedge b) = a \wedge b$, $D_{ca}(a \wedge b) = c \wedge b$, etc.

c) L'identité de Capelli.

Voici l'outil principal de ce paragraphe :

Théorème 2.5.
Soit $f \in \mathbf{R}[a,b,c]$ un polynôme. On a l'identité :

$$D_{yy} \circ D_{xx}(f) + D_{xx}(f) - D_{xy} \circ D_{yx}(f) = (x \wedge y)\Omega_{xy}(f)$$

où l'on a posé $\Omega_{xy}(f) = \dfrac{\partial^2 f}{\partial x_1 \partial y_2} - \dfrac{\partial^2 f}{\partial y_1 \partial x_2}$.

Démonstration. Comme les opérations de polarisation sont linéaires, il suffit de montrer la formule pour les monômes de la forme $f = x_1^{\alpha_1} x_2^{\alpha_2} y_1^{\beta_1} y_2^{\beta_2}$. C'est un calcul sans malice, on a :

$$\begin{aligned} D_{xy} \circ D_{yx} f = \big[\alpha_1(\beta_1+1) + \alpha_2(\beta_2+1)\big] f &+ \alpha_1\beta_2 x_1^{\alpha_1-1} x_2^{\alpha_2+1} y_1^{\beta_1+1} y_2^{\beta_2-1} \\ &+ \alpha_2\beta_1 x_1^{\alpha_1+1} x_2^{\alpha_2-1} y_1^{\beta_1-1} y_2^{\beta_2+1}, \end{aligned}$$

puis $D_{yy}f = (\beta_1 + \beta_2)\, f$ et $D_{xx}f = (\alpha_1 + \alpha_2)\, f$, d'où

$$D_{yy} \circ D_{xx}(f) + D_{xx}(f) = (\beta_1 + \beta_2 + 1)(\alpha_1 + \alpha_2)\, f.$$

Enfin on a $\Omega_{xy}f = \alpha_1\beta_2 x_1^{\alpha_1-1}x_2^{\alpha_2}y_1^{\beta_1}y_2^{\beta_2-1} - \alpha_2\beta_1 x_1^{\alpha_1}x_2^{\alpha_2-1}y_1^{\beta_1-1}y_2^{\beta_2}$, d'où le résultat en multipliant par $x \wedge y$.

3. Calcul des invariants sous $SL(2)$.

a) Détermination des invariants.

Le théorème essentiel de ce paragraphe est le suivant :

Théorème 3.1.
Soit f un polynôme de $\mathbf{R}[a, b, c]$ invariant sous $SL(2)$. Alors f est un polynôme en $b \wedge c$, $c \wedge a$, $a \wedge b$.

Démonstration. Dans ce qui suit on désigne par R le sous-anneau de $\mathbf{R}[a, b, c]$ engendré par $b \wedge c$, $c \wedge a$ et $a \wedge b$. En vertu de 1.4 on peut supposer que f est homogène (séparément) en a, b, c. On ordonne ces polynômes de la façon suivante : si le degré total de f est plus petit que celui de g on dit qu'on a $f \leq g$. Si ces polynômes ont même degré total on ordonne suivant l'ordre lexicographique des degrés en a, b, c (le degré en a prédomine). On raisonne alors par l'absurde en supposant qu'il y ait un polynôme invariant trihomogène f qui n'est pas dans R et on choisit un tel polynôme minimal pour l'ordre ci-dessus. Ce polynôme n'est pas une constante (sinon il serait dans R). Supposons qu'il soit de degré $p > 0$ en a et q en b (la démonstration est la même, *mutatis mutandis* dans les autres cas). On applique l'identité de Capelli à f avec $x = a$ et $y = b$. On a donc

$$D_{bb} \circ D_{aa}(f) + D_{aa}(f) - D_{ab} \circ D_{ba}(f) = (a \wedge b)\Omega_{ab}(f).$$

Comme f est invariant il en est de même des polarisations de f (cf. 2.3). Il en résulte que $(a \wedge b)\Omega_{ab}(f)$ est invariant et, comme $a \wedge b$ l'est aussi, $\Omega_{ab}(f)$ est un invariant. Comme $\Omega_{ab}(f)$ est de degré total strictement plus petit que celui de f, l'hypothèse de minimalité montre que $\Omega_{ab}(f)$ est dans R. D'autre part, $D_{ba}f$ est un polynôme invariant de même degré total que f mais de degré plus petit en a (cf. 2.2.b). Vu l'hypothèse de minimalité du choix de f, il est donc dans R. Mais alors, la stabilité de R par polarisation montre que $D_{ab} \circ D_{ba}(f)$ est aussi dans R. On en déduit que $D_{bb} \circ D_{aa}(f) + D_{aa}(f)$ est dans R, mais ce polynôme n'est autre que $(q + 1)pf$, et comme $(q + 1)p$ est non nul, f est dans R, ce qui est absurde.

b) Détermination des relations.

Comme on l'a annoncé, il n'y en a aucune :

Théorème 3.2.
Le sous-anneau R de $\mathbf{R}[a, b, c]$ engendré par $b \wedge c$, $c \wedge a$ et $a \wedge b$ est isomorphe à $\mathbf{R}[X, Y, Z]$ autrement dit, il n'y a pas de polynôme $P(X, Y, Z)$ non nul qui vérifie $P(b \wedge c, c \wedge a, a \wedge b) = 0$.

Démonstration. On montre d'abord le lemme suivant :

Lemme 3.3.
Soient x, y, z trois réels > 0. Il existe des réels $a_1, a_2, b_1, b_2, c_1, c_2$ tels que l'on ait $x = b \wedge c$, $y = c \wedge a$ et $z = a \wedge b$.

Démonstration. Il suffit de prendre par exemple $a_1 = 1$, $a_2 = 0$, $b_1 = -\frac{x}{y}$, $b_2 = z$, $c_1 = 0$, $c_2 = -y$.

Revenons à 3.2 en supposant qu'un polynôme $P(X, Y, Z)$ s'annule sur les polynômes $b \wedge c$, $c \wedge a$ et $a \wedge b$. En vertu de 3.3 ce polynôme serait nul sur $(\mathbf{R}^{+*})^3$ donc serait identiquement nul.

4. Le cas de $GL(2)$.

a) Invariants absolus et relatifs.

Proposition 4.1.
Les seuls polynômes de $\mathbf{R}[a, b, c]$ invariants sous $GL(2)$ sont les constantes.

Démonstration. Soit $P \in \mathbf{R}[a, b, c]$ un polynôme homogène de degré d et soit $h \in GL(2)$ l'homothétie de rapport $\lambda \neq \pm 1$. On a $P.h = \lambda^d P$, de sorte que P n'est invariant par h que si d est nul.

Rappelons qu'un polynôme $P \in \mathbf{R}[a, b, c]$ est un invariant relatif sous $GL(2)$ si pour tout $u \in GL(2)$ il existe $\lambda(u) \in \mathbf{R}^*$ tel que $P.u = \lambda(u)P$.

Proposition 4.2.
Les invariants relatifs de $\mathbf{R}[a, b, c]$ sous $GL(2)$ sont des polynômes homogènes en a, b, c.

Démonstration. Si P est un tel polynôme, écrit comme somme de polynômes homogènes $P = P_0 + \cdots + P_d$ et si on applique l'homothétie h de rapport $\lambda \neq \pm 1$ on a $P.h = P_0 + \lambda P_1 + \cdots + \lambda^d P_d = \lambda(h)P$ et cela impose que P est homogène.

La proposition suivante précise le multiplicateur $\lambda(u)$:

Proposition 4.3.

1) L'application $u \mapsto \lambda(u)$ est un homomorphisme de groupes de $GL(2)$ dans $\mathbf{R}^*$.
2) Le multiplicateur $\lambda(u)$ est une fonction polynomiale des coefficients de la matrice u.
3) Il existe un entier $n \in \mathbf{N}$ tel que $\lambda(u) = (\text{dét}\, u)^n$.

Démonstration. La première assertion est immédiate et la seconde résulte d'un petit calcul d'identification. Pour la troisième, on considère la matrice adjointe u^* (matrice des cofacteurs), de sorte qu'on a $uu^* = (\text{dét}\, u)\,\text{Id}$. Si P est un polynôme homogène de degré d, on a donc $P.uu^* = (\text{dét}\, u)^d P = \lambda(u)\lambda(u^*)P$, de sorte que le polynôme $\lambda(u)$ en les coefficients $\alpha, \beta, \gamma, \delta$ de la matrice u divise le polynôme $(\alpha\delta - \beta\gamma)^d$. Mais, comme $\alpha\delta - \beta\gamma$ est irréductible comme polynôme en $\alpha, \beta, \gamma, \delta$, $\lambda(u)$ est bien une puissance de $\text{dét}\, u = \alpha\delta - \beta\gamma$.

Corollaire 4.4.
Les invariants relatifs de $\mathbf{R}[a, b, c]$ sous $GL(2)$ sont les polynômes homogènes en

$b \wedge c$, $c \wedge a$ et $a \wedge b$. Pour un tel polynôme homogène de degré d le multiplicateur est dét u^d.

Démonstration. En vertu de 4.3.3 les invariants relatifs de $GL(2)$ sont des invariants absolus de $SL(2)$ donc des polynômes en les invariants élémentaires.

Proposition 4.5.
Les fractions rationnelles invariantes sous $GL(2)$ sont de la forme $\frac{F}{G}$ où F et G sont des polynômes en $b \wedge c$, $c \wedge a$ et $a \wedge b$, homogènes de même degré.

Démonstration. Soit $\frac{F}{G}$ une fraction rationnelle invariante, avec F, G premiers entre eux. On a donc $\frac{F}{G} = \frac{F.u}{G.u}$, d'où $F\,(G.u) = G\,(F.u)$. Comme F divise le second membre et est premier avec G il divise $F.u$. Comme les degrés de F et $F.u$ sont égaux on a $F.u = \lambda(u)F$ où $\lambda(u)$ est un scalaire. Autrement dit F (et de même G) est un invariant relatif. On conclut avec 4.3.

Annexe 2 : Le théorème de Bolyai

Nous montrons dans cette annexe que, pour calculer les aires des polygones, la méthode de découpage et recollement, que l'on utilise dès l'école primaire, est toujours applicable.

1. Énoncé.

Nous appellerons polygone une réunion finie de polygones convexes. On dit qu'une réunion de polygones A_i $(i = 1, \cdots, n)$ est **presque disjointe** si les intersections $A_i \cap A_j$ pour $i \neq j$ sont des réunions finies de segments de droites.
On dit que deux polygones A et B sont **équivalents** (sous-entendu par découpage et recollement) et on note $A \sim B$ si on peut écrire

$$A = \bigcup_{i=1}^{n} A_i, \qquad B = \bigcup_{i=1}^{n} B_i,$$

où les A_i, B_i sont des polygones, où les réunions sont presque disjointes et où, pour tout i, A_i et B_i sont directement isométriques (i.e. il existe un déplacement g_i de $\mathbf{R}^2$ tel que $g_i(A_i) = B_i$). Cela signifie exactement qu'on a découpé A en les A_i, qu'on a déplacé ceux-ci (sans les retourner, cf. 4.2) pour obtenir les B_i et qu'on a recollé les B_i pour obtenir B, le tout, bien entendu, sans perte ni chevauchement (sauf éventuellement sur les bords). C'est bien la notion usuelle.
Si G est un sous-groupe du groupe des déplacements, nous préciserons parfois que les polygones sont G-équivalents pour signifier que les g_i doivent être pris dans G. On note $\sim_G$ la relation correspondante. Dans ce qui suit nous utiliserons cette notion en prenant pour G le groupe S engendré par les symétries centrales (c'est-à-dire le groupe dont les éléments sont les translations et les symétries centrales).

Le lemme suivant est laissé au lecteur :

Lemme 1.1.
a) Deux polygones directement isométriques (resp. G-isométriques) sont équivalents (resp. G-équivalents).
b) La relation $\sim$ (resp. $\sim_G$) est une relation d'équivalence.

Il résulte des axiomes des aires et du fait qu'un segment est d'aire nulle que si A et B sont équivalents ils ont même aire. La question intéressante est la réciproque :

Théorème 1.2.
Soient A et B deux polygones de même aire. Alors A et B sont équivalents par découpage et recollement (Bolyai, 1832). On peut même préciser (Hadwiger-Glur, 1951) qu'ils sont S-équivalents.

Nous nous contenterons ici de montrer le théorème de Bolyai. La variante d'Hadwiger-Glur est plus difficile, même si elle n'utilise que des arguments élémentaires. Le lecteur en trouvera une preuve dans [B1] ou [B2]. On montre que tout polygone est S-équivalent à un rectangle (un peu comme en 2.1 ci-dessous), puis, et c'est le point délicat, que deux rectangles de même aire sont S-équivalents.

2. Le lemme fondamental.

Nous allons montrer le lemme suivant :

Lemme 2.1.
Soit A un polygone quelconque. Il existe un rectangle R équivalent à A dont un côté est de longueur unité.

Montrons d'abord que ce lemme implique le théorème de Bolyai. En effet, soient A et B des polygones de même aire. D'après le lemme, il existe des rectangles R_A et R_B, dont un côté est l'unité, et respectivement équivalents à A et B. Comme l'aire est conservée par équivalence, les rectangles R_A et R_B ont même aire, donc leurs deuxièmes côtés sont égaux. Mais alors on passe de l'un à l'autre par un déplacement (mais pas nécessairement un déplacement de S !), donc ces rectangles sont équivalents et on a $A \sim R_A \sim R_B \sim B$ donc $A \sim B$ par transitivité.

3. Preuve du lemme fondamental.

Notons d'abord qu'il suffit d'établir le lemme 2.1 pour un triangle. En effet, si A est un polygone quelconque, on peut toujours l'écrire comme réunion presque disjointe de triangles $T_1, T_2, \cdots, T_n$. On peut alors trouver des rectangles R_i dont un côté est 1 et l'autre l_i et qui sont équivalents aux T_i. On obtient un rectangle R de côté $l_1 + \cdots + l_n$ équivalent à A en recollant les R_i le long de leurs côtés de longueur 1.

Le pas suivant est de découper un triangle pour en faire un parallélogramme :

Lemme 3.1.
Soit T un triangle. Il existe un parallélogramme P tel que $T \sim P$.

Démonstration. Soit ABC le triangle et M et N les milieux de $[AB]$ et $[AC]$. On sait que (MN) est parallèle à (BC). Considérons la symétrie σ de centre N. On a $\sigma(A) = C$. Notons D le symétrique de M par rapport à N (cf. figure 6). Alors,

la droite (CD) est image de (AM) par σ, donc elle lui est parallèle. Il en résulte que le quadrilatère $MDCB$ est un parallélogramme. Le triangle T est réunion du triangle AMN et du trapèze $MNCB$, tandis que le parallélogramme $MDCB$ est réunion du même trapèze et du triangle symétrique CDN. On a bien montré que le triangle et le parallélogramme sont équivalents.

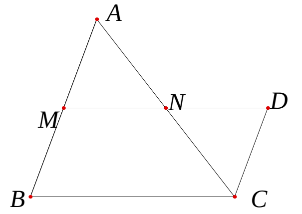


Figure 6

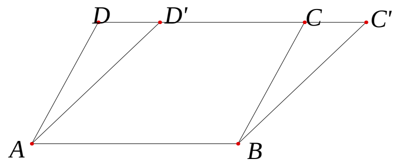


Figure 7

On passe maintenant d'un parallélogramme à un autre, mais dont un côté est 1 :

Lemme 3.2.
Soit P un parallélogramme.
a) Il existe un parallélogramme P', dont un côté est de longueur rationnelle, tel que $P \sim P'$.
b) Il existe un parallélogramme P'', dont un côté est de longueur p, avec $p \in \mathbf{N}^*$, tel que $P \sim P''$.
c) Il existe un parallélogramme P''' dont un côté vaut 1 et tel que $P \sim P'''$.

Démonstration. a) Soit $ABCD$ le parallélogramme donné et soit $D' \in [DC]$ un point tel que la longueur AD' soit un nombre rationnel. Un tel point existe car lorsque D' varie dans $[DC]$ la longueur AD' varie dans un intervalle de $\mathbf{R}$, non réduit à un point (car un segment n'est pas un arc de cercle !), donc cet intervalle contient des rationnels.
On considère alors le point C' de (DC) tel que $\overrightarrow{DD'} = \overrightarrow{CC'}$, cf. figure 7. Le quadrilatère $ABC'D'$ est un parallélogramme. De plus, on passe de $ABCD$ à $ABC'D'$ par découpage du triangle ADD' que l'on recolle, après translation de vecteur $\overrightarrow{AB}$, en BCC'. Enfin, le parallélogramme $ABC'D'$ a un côté rationnel, à savoir AD'.

b) Grâce à a) on peut partir d'un parallélogramme $ABCD$ dont le côté AD est rationnel. On pose $AD = \dfrac{p}{q}$ et on partage le côté AB en q parties égales, par des points $A_0 = A, A_1, \cdots, A_q = B$. On mène alors, par les A_i, les droites parallèles à (AD) qui coupent (CD) en des points C_i. Ces droites découpent P en q parallélogrammes de base AB/q. On transporte alors ces parallélogrammes et on les empile sur AA_1C_1D, cf. figure 8. On obtient bien ainsi un parallélogramme dont un côté vaut $qAD = p \in \mathbf{N}$.

c) On dispose maintenant d'un parallélogramme avec un côté entier, disons $AD = p$. On partage alors $[AD]$ en p parties égales, toutes de longueur 1 par des points $A_0 = A, A_1, \cdots, A_p = D$ et on mène les parallèles à (AB) par les A_i, qui recoupent (BC) en les C_i. On obtient ainsi p petits parallélogrammes que l'on découpe et que l'on transporte pour les mettre, en ligne, à côté de ABC_1A_1, cf. figure 9. On a bien maintenant un parallélogramme dont un côté vaut 1.

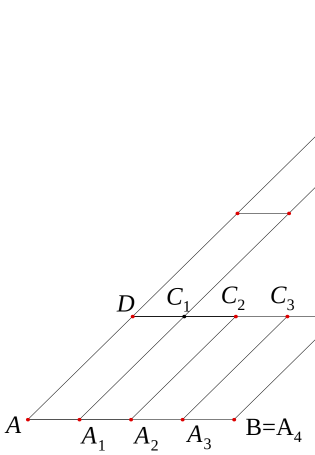


Figure 8

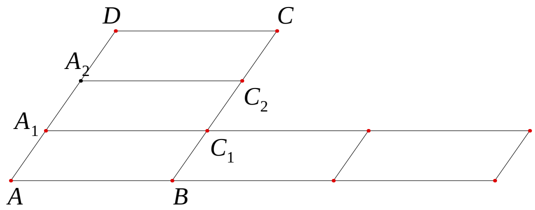


Figure 9

Pour en finir avec le lemme 2.1 il ne reste plus qu'à montrer le résultat suivant :

Lemme 3.3.
Soit $P = ABCD$ un parallélogramme dont le côté AB vaut 1. Alors P est équivalent à un rectangle de côté 1.

Démonstration. On suppose d'abord que le projeté C' de B sur la droite (CD) est dans le segment $[CD]$. On considère alors le point D' tel que $\overrightarrow{CC'} = \overrightarrow{DD'}$. Le quadrilatère $ABC'D'$ est un rectangle qui s'obtient à partir de $ABCD$ par découpage de BCC' que l'on recolle en ADD'. Comme il a un côté égal à $AB = 1$ on a gagné. Le raisonnement est identique si le projeté A' de A sur (DC) est dans le segment $[DC]$.
Lorsque les deux projetés A' et C' sont hors de $[CD]$, on partage AD en n parties égales par des points A_i et on mène les parallèles à (AB) par les A_i. Elles recoupent (BC) en C_i. On choisit n assez grand pour que le projeté de A ou de B sur (A_1C_1) soit dans $[A_1C_1]$. La même propriété vaut aussi pour tous les petits parallélogrammes empilés sur ABC_1A_1. On peut donc, par le premier cas, découper ces parallélogrammes pour en faire des petits rectangles de côté 1, cf. figure 10. Il ne reste plus qu'à empiler les rectangles obtenus et on a gagné.

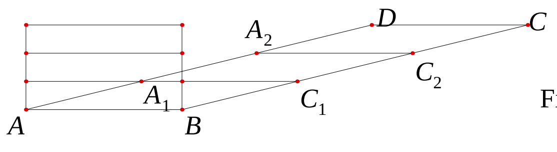


Figure 10

Remarques 3.4.
1) On peut montrer un théorème analogue avec des unions disjointes au lieu de presque disjointes, voir le joli livre de Marc Guinot [G].
2) On notera que la démonstration ci-dessus fait un usage essentiel de l'axiome d'Archimède. On peut montrer que le théorème n'est pas vrai sur un corps non archimédien.

4. Quelques problèmes.

Le théorème de Bolyai assure donc qu'on peut toujours trouver un puzzle pour passer d'un polygone à un polygone de même aire (et il fournit un algorithme). Une question souvent plus amusante est de trouver le puzzle le plus simple possible. Voici deux exemples de tels problèmes :

1) On se donne un carré. Produire un puzzle pour le transformer en deux carrés, en trois carrés.

2) Il était une fois un pâtissier, féru de géométrie, mais distrait. Il possédait un moule à tarte en forme de triangle scalène (c'est-à-dire un triangle quelconque donc non isocèle). Pour faire sa tarte il en mesura les trois côtés, réalisa la pâte aux mesures et la couvrit de fruits. Hélas, il s'aperçut alors que le triangle qu'il venait de réaliser n'était pas le bon, mais son symétrique par rapport à une droite. Sauf à faire une tarte à l'envers il ne pouvait donc mettre sa tarte dans le moule. Comme il connaissait le théorème de Bolyai il se mit en quête d'un découpage lui permettant de rattraper sa bévue. Pouvez-vous l'aider ?